CS Colloquium 04 March, 16:20-17:40 «Mirror-Prox Algorithm with Linear Convergence Rate and its Application for Dynamic Loss Scaling»



Aleksandr Beznosikov MIPT, ISP RAS, HSE University

During this talk, we extend the classical machine learning problem of the empirical risk minimization by adding weights to the loss function of each training object. We select these weights adaptively during the optimization process by modifying the original minimization problem to a min-max formulation. The maximization is performed over the above weights. To solve this saddle-point problem, we explore different variations of the Mirror-Prox algorithm: classical, optimistic, stochastic and with variance reduction. For all these methods, we obtain the convergence rates for the original min- max problem. Based on these algorithms we propose a novel deep learning (DL) optimizer ALSO – Adaptive Loss Scaling Optimizer. To justify our approach, we consider a wide range of DL problems, from Tabular DL to image generation tasks.





